Reduced levels of microsatellite variability on the neo-Y chromosome of *Drosophila miranda*

Doris Bachtrog and Brian Charlesworth

Background: In many species, sex is determined by a system involving X and Y chromosomes, the latter having lost much of their genetic activity. Sex chromosomes have evolved independently many times, and several different mechanisms responsible for the degeneration of the Y chromosome have been proposed. Here, we have taken advantage of the secondary sex chromosome pair in *Drosophila miranda* to test for the effects of evolutionary forces involved in the early stages of Y-chromosome degeneration. Because of a fusion of one of the autosomes to the Y chromosome, a neo-Y chromosome and a neo-X chromosome have been formed, resulting in the transmission of formerly autosomal genes in association with the sex chromosomes.

Results: We found a 25-fold lower level of variation at microsatellites located on the neo-Y chromosome compared with homologous loci on the neo-X chromosome, or with autosomal and X-linked microsatellites. Sequence analyses of the region flanking the microsatellites suggested that the neo-sex chromosomes originated about 1 million years ago.

Conclusions: Variability of the neo-Y chromosome of *D. miranda* is substantially reduced below expectations at mutation-drift equilibrium. Such a reduction is predicted by theories of the degeneration of the Y chromosome. Another possibility is that there is little or no mutation at microsatellite loci on a non-recombining chromosome such as the neo-Y, but this seems inconsistent with other data.

Address: Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK.

Correspondence: Doris Bachtrog
E-mail: Doris.Bachtrog@ed.ac.uk

Received: **17 May 2000**
Revised: **12 July 2000**
Accepted: **12 July 2000**

Published: **9 August 2000**

Current Biology 2000, 10:1025–1031

0960-9822/00/$ – see front matter

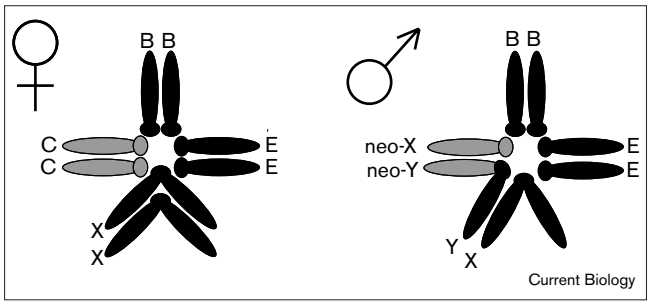
Background

The evolution of separate sexes is often associated with the evolution of morphologically distinct sex chromosomes [1]. The most familiar form of genetic sex determination involves homogametic females (XX) and heterogametic males (XY). It is generally believed that the X and Y chromosomes have descended from initially homologous chromosomes [1,2]. Suppression of recombination between the ancestral proto-sex chromosomes has led to the evolution of morphologically and functionally distinct sex chromosomes [3,4], with loss of genetic activity from the Y chromosome. In response, the X chromosome often becomes dosage-compensated [1,3]. Various population genetic models have been proposed to account for the degeneration of the Y chromosome. Most of these postulate the accumulation of deleterious mutations on the Y chromosome, driven by its lack of recombination [3,5,6]. A feature that is common to all these processes is that the non-recombining Y chromosome has a much smaller effective population size, N_e, than expected under neutral conditions [3,7,8].

As the Y chromosomes of species such as humans or *Drosophila melanogaster* are very ancient and almost completely genetically eroded, they are not suitable for examining the processes involved in the early stages of degeneration [3]. The newly formed secondary sex chromosomes of *D. miranda* provide an excellent opportunity for studying these processes [9]. As a result of a fusion of an autosome (Muller's element C [10]) to the Y chromosome (Figure 1), formerly autosomal genes are now transmitted solely through males. As male *Drosophila* lack crossing over [11], this newly formed 'neo-Y' chromosome is sheltered from recombination and should therefore be subject to the same evolutionary forces that led to the degeneration of the true Y chromosome. In fact, the neo-Y and the homologous unfused 'neo-X' (see Figure 1) are both evolving altered chromosomal structures. The neo-Y chromosome is well on its way to degeneration [12,13], while the neo-X chromosome is becoming dosage-compensated [14–16].

As the long-term equilibrium level of genetic diversity at a locus is proportional to N_e [17], the hypothesis that the evolutionary degeneration of the neo-Y chromosome of *D. miranda* is associated with a reduction in N_e can be tested by comparing variability at loci on the neo-Y with that at homologous loci on the neo-X chromosome [18,19]. To obtain an insight into the genetic diversity of the neo-Y chromosome, we characterised variability at nine microsatellites located on the neo-sex chromosomes of *D. miranda*. For comparison, polymorphism at nine X-linked

Figure 1



Karyotype of *D. miranda*. Females of *D. miranda* have 2n = 10 chromosomes, whereas males have an odd number (2n = 9) because of the fusion of chromosome 3 (Muller's element C) with the Y chromosome (dot chromosomes not shown). This fused chromosome is called the neo-Y chromosome. Its homologue, which is present in two copies in females and only one copy in males (a characteristic of X chromosomes), is the neo-X chromosome.

and nine autosomal microsatellites was also investigated. Sequence analyses of the region flanking the microsatellite loci were also conducted, confirming the presence of the microsatellites on the neo-sex chromosomes and allowing us to estimate the age of this chromosomal system. Implications of our results for theories of the degeneration of the Y chromosome are discussed.

Results

Microsatellite variability of the neo-X versus the neo-Y chromosome

Levels of variability across the different loci are tabulated in Table 1. The microsatellites located on the neo-X chromosome harboured substantial variation, with a mean heterozygosity ($\bar{H}$) of 0.57 and an average variance in repeat number ($\bar{V}$) of 6.08. In contrast, the homologous neo-Y-linked copies almost completely lacked variability: only a single segregating variant was detected ($\bar{H}$ = 0.02 and $\bar{V}$ = 0.19). Two loci that yielded a neo-X-linked copy failed to amplify a neo-Y-linked allele (DPS3003 and DPS3004), indicating either a mutation in the priming site or a deletion of the neo-Y-linked microsatellite. In all other cases, sequencing verified that a microsatellite structure was present on the neo-Y chromosome.

A direct comparison of the neo-X and the neo-Y microsatellites revealed an approximately 25-fold difference in variability. Under random mating and a 1:1 sex ratio, the effective population size of the neo-Y is one-third that of the neo-X, implying that its diversity is expected to be lower by the same factor, assuming similar mutation rates [17,20]. Correcting for the difference in N_e by dividing the variability of the neo-X by three still gave an eightfold higher variability of microsatellites on the neo-X chromosome, suggesting that the difference in chromosome copy number cannot account for the observed reduction in variability on the neo-Y chromosome.

X-chromosomal and autosomal microsatellites

For comparison, we also analysed variability at several X-linked and autosomal microsatellites (Table 2). Neo-X-linked, X-chromosomal and autosomal microsatellites showed similar levels of mean heterozygosity ($\bar{H}$ = 0.57, $\bar{H}$ = 0.49 and $\bar{H}$ = 0.46, without correction for differences in N_e) and average variance in repeat number ($\bar{V}$ = 6.08, $\bar{V}$ = 6.86 and $\bar{V}$ = 3.74). This suggests that mutation rates of the neo-X-linked microsatellites are not unusual compared with other loci.

Population differentiation

A study of the *period* locus using the same lines of *D. miranda* has suggested that the Canadian and American populations are genetically distinct [18]. To examine this possibility with our data, F_{ST} values, which measure the extent of differentiation between populations [17], and an exact contingency-table test for population subdivision

Table 1

Variability of microsatellites located on the neo-sex chromosomes of *D. miranda*.

	Neo-X chromosome			Neo-Y chromosome		
Locus	Heterozygosity	Variance	Alleles	Heterozygosity	Variance	Alleles
DPS3001	0.15	2.08	2	0.00	0.00	1
DPS3002	0.83	13.84	7	0.00	0.00	1
DPS3003	0.78	23.88	6	–	–	–
DPS3004	0.60	4.09	4	–	–	–
DPS3005	0.60	1.91	4	0.00	0.00	1
DPS3006	0.65	4.06	4	0.00	0.00	1
DPS3008	0.68	4.24	5	0.00	0.00	1
engrailed	0.28	0.15	2	0.15	1.33	2
129*	0.57	0.45	3	0.00	0.00	1
Average	0.57	6.08	4.11	0.02	0.19	1.14

*Locus 129 was kindly provided by M. Noor.

Table 2

Variability of X-linked and autosomal microsatellites in *D. miranda*.

	Locus	Heterozygosity	Variance	Alleles
X chromosome	DPSX001	0.81	25.66	6
	DPSX002	0.00	0.00	1
	DPSX003	0.69	16.97	6
	DPSX004	0.51	4.70	4
	DPSX006	0.51	2.02	4
	DPSX007	0.72	5.30	5
	DPSX008	0.58	0.93	4
	DPSX009	0.57	6.21	3
	DPSX010	0.00	0.00	1
	Average	0.49	6.86	3.78
Chromosome 2	DPS2001	0.15	0.08	2
	DPS2002	0.57	0.45	3
	DPS2003	0.28	0.15	2
	DPS2004	0.64	1.24	4
	DPS2005	0.61	4.27	5
	DPS2006	0.83	7.70	8
	DPS2007	0.53	19.36	4
	Mlc*	0.00	0.00	1
	trop1*	0.49	0.45	3
	Average	0.46	3.74	3.56

*Mlc and trop1 are microsatellite loci [48].

were calculated over all microsatellite loci. Neither showed significant differentiation among populations. The mean of F_{ST} over all loci was 0.011, which is not significantly different from zero, using a permutation test to estimate significance levels ($p = 0.12$). This indicates that the subdivision seen at *period* is specific to this particular locus, perhaps because of local adaptation [18]. Sequence variation at five additional loci in *D. miranda* also yielded no indication of geographic structure [21]. This lack of population subdivision justifies the use of coalescent simulations that assume panmixia, in order to estimate the reduction in N_e of the neo-Y chromosome.

Estimates of the reduction in the effective population size of the neo-Y using coalescent process simulations

As described in the Materials and methods, two features of the data were included in our simulations. First, the loci scored on the neo-X and neo-Y were homologous (they should have the same mutation rate), while mutation rates among loci could vary substantially. Second, the microsatellites on the neo-Y share one genealogy, whereas the genealogies for the neo-X-linked loci are probably completely independent. To account for this difference, one tree with seven completely linked loci was generated for the neo-Y chromosome, whereas seven independent trees were computed for the neo-X chromosome. For generating mutations, estimates of $\theta = 4N_e\mu$, where μ is the mutation rate, were drawn from a gamma distribution, using the same values for homologous loci on the two chromosomes. After each simulation, the average variance in repeat number over all loci was calculated for the neo-X ($\bar{V}_X$) and neo-Y ($\bar{V}_Y$) chromosomes.

To determine whether microsatellites on the neo-Y chromosome show significantly reduced diversity, the proportion of runs (P_c) in which the simulated value of $\Delta V = (\bar{V}_X - \bar{V}_Y)/\bar{V}_X$ was equal to or larger than the observed value of ΔV was determined. The values of θ were multiplied by a factor k before mutations were laid down on the neo-Y tree, to account for its possibly lower population size. If $P_c \leq 0.05$, the parameters of the simulations are incompatible with the data at the 5% probability level. Under neutrality and a Poisson distribution of family size for males and females, k should be 0.33, because of differences in chromosome copy number [22]. Using a value of $k = 0.33$ is inconsistent with our data; only 2.0% of the runs yielded a ΔV equal or larger than the observed quantity (Table 3). With $k = 0.22$, about 5% of the simulations were compatible with the observed difference in microsatellite variability, giving an upper bound of a fivefold reduction in N_e for the neo-Y chromosome compared with the neo-X chromosome.

Divergence data for the sequences flanking the microsatellites

One neo-X- and one neo-Y-linked allele from each locus were sequenced to confirm the presence of the microsatellites on the neo-sex chromosomes and allow us to estimate the age of this chromosomal system. Among the 733 basepairs of flanking sequence determined, 18 sites differed between the neo-X and the neo-Y chromosome, translating into a Jukes–Cantor distance of 0.025 per site (excluding indel polymorphism). Correcting for ancestral polymorphism gave a net divergence between the two chromosomes of about 0.021 (the average synonymous site diversity for two autosomal genes has been estimated to be 0.004 [21]). Assuming a substitution rate for noncoding sequences of 1.2×10^{-8} per site per year for *Drosophila* [23], this yielded a total divergence time of 1.72 million years, dating the origin of the neo-sex chromosome system to about 0.86 million years ago (95% confidence interval = 0.72–1.00 million years). Sequence divergence of the *Lcp* genes, which are also located on the neo-sex chromosomes of *D. miranda*, has indicated a

Table 3

Results of coalescence simulations.

	$\bar{V}_X$	$\bar{V}_Y$	Proportion of replicates with ΔV simulated > ΔV observed
$k = 1$	5.50	5.49	0.07%
$k = 0.33$	5.49	1.83	2.0%
$k = 0.22$	5.50	1.21	5.6%
$k = 0.11$	5.51	0.61	21.2%
$k = 0.05$	5.50	0.27	54.1%

slightly older age of this chromosomal system (about 1.25 million years [18]).

Discussion

Neo-X versus neo-Y

Analyses of microsatellites located on the neo-sex chromosomes of *D. miranda* reveal that the variability of the neo-Y chromosome is greatly reduced. In contrast, levels of polymorphism at Y-linked microsatellites in humans, corrected for differences in N_e, are similar to those of autosomal loci [24–26]. This is consistent with the human Y chromosome being much older and having lost most of its genetic activity [27]. A search for functional genes on the non-recombining region of the Y in humans has revealed a total of only 20 loci [28]. With many fewer targets for selection, forces which cause the degeneration of the Y chromosome in association with the reduction in N_e should no longer be operating [3]. Nevertheless, nucleotide polymorphism at the Y-linked *Dhc-Yh3* gene in *D. melanogaster* and *D. simulans* is significantly reduced compared with X-linked and autosomal loci, even after correction for differences in N_e, which was attributed to either strong sexual selection in males or to adaptive hitchhiking [29].

Sequence data on the *Lcp* gene family, which is also located on the neo-sex chromosomes in *D. miranda*, showed only an about threefold reduction in polymorphism on the neo-Y compared to the neo-X chromosome [18]. The confidence interval for this ratio is, however, large because of the low level of polymorphism observed for both chromosomes. In addition, comparisons with an X-linked locus revealed that variability of the neo-X-linked copies of the *Lcp* genes also appears to be reduced, which may reflect selection for *cis*-acting sites used in dosage compensation on this newly evolving X chromosome [18].

Two other studies on the early stages of the degeneration of Y chromosomes using DNA sequence data have been performed. A Y-linked gene in the plant *Silene latifolia* displayed a similar magnitude of reduction in variability [19]. Nucleotide polymorphism at a locus on the neo-Y chromosome of *D. americana* suggests only a modest reduction in the effective population size [30], consistent with the likelihood of ongoing recombination in this chromosomal system [30] and the lack of degeneration of the neo-Y in this species [31].

Only a single segregating variant was observed on the neo-Y chromosome of *D. miranda*, whereas variability at homologous loci on the neo-X chromosome was comparable with microsatellites on the X chromosome or autosomes. There are two possible causes for the lower than expected level of polymorphism: microsatellites might either show a decreased mutation rate on the neo-Y chromosome, or the effective population size of the neo-Y chromosome is reduced below simple expectation.

If recombination enhances microsatellite mutability, loci located on the neo-Y chromosome should harbour less polymorphism. There is, however, no evidence for a role of crossing over in generating mutations at microsatellite loci. Studies in recombination-deficient *Escherichia coli* or yeast did not detect any change of the mutational behaviour of microsatellites in these strains [32,33]. No direct comparisons of mutation rates in *Drosophila* between autosomal and Y-linked microsatellites have been carried out. In humans, however, microsatellites located on the non-recombining region of the Y chromosome have comparable mutation rates to autosomal loci [34,35], further indicating that the general mutational mechanism of microsatellites is independent of recombination. Another concern from the use of microsatellites is the possibility of allele-specific mutation rates, with longer alleles having higher mutation rates [36]. Subtracting the flanking sequence from the length of the PCR products scored here, the mean population repeat number of loci on the neo-X chromosome was estimated as 9.15, whereas microsatellites on the neo-Y chromosome showed an average of 9.24 repeats. If mutation rates were higher for longer alleles, then the neo-Y-linked loci would be expected to be slightly more mutable. Although we cannot completely rule out the possibility of reduced mutation rates of the neo-Y-linked loci, this is not supported by current data on the mutational process at microsatellite loci. Further experiments, especially on mutation rates for chromosomes propagated only through male *Drosophila*, are needed to resolve this question.

The second explanation for the low variability observed is that the effective population size of the neo-Y chromosome is reduced below the neutral expectation of one-third of N_e for the neo-X chromosome. Because sex chromosomes and autosomes show differences in their sex-limited transmission, comparisons between them will reflect the effects on variability of sex-specific life history traits. As the Y chromosome is transmitted through males only, it is affected only by male life history, whereas X-linked genes are biased towards the female contribution, and autosomal genes are averaged between the two sexes. Sexual selection on males can at most reduce Y-linked variability to one-ninth that of the X [37], whereas we observed a 25-fold reduction. In our simulations, however, we cannot reject a reduction in N_e to one-ninth the neo-X value ($k = 0.11$, $P_c = 0.21$, Table 3). But the age-structure of *Drosophila* populations greatly reduces the effect of sexual selection below this maximum value (B.C., unpublished data; [38]). Using life history data from *D. melanogaster*, the likely ratio of Y to X-chromosomal diversity is ~0.19, and variability on the X is expected to be ~90% of that of the autosomes (B.C., unpublished data), consistent with observed levels of DNA sequence variability in *D. miranda* [21]. The value of $k = 0.19$ is close to the upper bound of the reduction inferred from the coalescent simulations. It is

thus just possible that sexual selection alone could explain the difference between neo-X and neo-Y chromosome variation. More extensive comparisons of X-linked and autosomal variability in this species should allow the quantification of the effect of sexual selection.

Alternatively, theories for explaining the degeneration of a non-recombining chromosome predict a reduction in N_e. The hitchhiking model of Y-chromosome degeneration postulates that the adaptive fixation of Y-linked mutations may drag along deleterious alleles present in the same chromosome, in which the advantageous mutation arose [8]. Such selective sweeps of beneficial mutations are known to substantially reduce neutral variability at linked sites [39,40]. The background selection model invokes the opposite form of hitchhiking: the selective elimination of strongly deleterious mutations in a genomic region with no recombination, such as the Y chromosome, implies that only chromosomes free of strongly deleterious mutations contribute to the ancestry of future generations [3]. This reduction in chromosome copy number accelerates the fixation of mildly deleterious mutations by random drift (a mutation is effectively neutral if $N_e s < 1$, where s is the selection coefficient) and reduces variability at linked neutral sites [41–43].

'Muller's ratchet' [44], the stochastic loss of the class of chromosomes carrying the smallest number of deleterious mutations, can also lead to the degeneration of a non-recombining Y chromosome [6]. In the absence of recombination and back mutation, this loss is irreversible, leading to a continuous increase in the mean number of deleterious mutations on the Y chromosome [45]. Under Muller's ratchet, all chromosomes in the population ultimately derive from the class with the fewest number of mutations [8,46], suggesting that the effective population size of the Y chromosome is probably considerably reduced below the actual number of chromosomes present. Interference between many weakly selected sites at a non-recombining chromosome (known as the 'Hill–Robertson effect') reduces the efficacy of selection [47]. This can cause a substantial reduction in fitness, as a result of an increased abundance of slightly deleterious variants on the Y [7], and is also associated with reduced diversity if there are many sites on the chromosome [7]. Given the evidence for substantial degeneration of the neo-Y chromosome of *D. miranda* [12,13], over a time scale of about 1 million years (see above and [18]), our evidence for a large reduction in variability is consistent with the operation of one or more of these processes.

Conclusions

Our analyses reveal that the variability of microsatellites on the neo-Y chromosome is substantially reduced, indicating a large reduction in its effective population size. This is consistent with mutation-driven degeneration of this chromosome, because of its lack of recombination. A lower mutation rate of the microsatellites on the neo-Y chromosome, which could in principle also explain the observed lack of variation, is not supported by other data. Further characterisation of the properties of molecular evolution and variation at loci on this chromosome should help to elucidate the processes involved in Y-chromosome degeneration.

Materials and methods

Strains

The following *D. miranda* lines were used for the microsatellite analyses, with their geographic origin given in brackets: 0101.3, 0101.4, 0101.5, 0101.7 (Port Coquitlam, British Columbia, Canada), 0101.9, MA28, MA32 (Mather, California), SP138, SP235, SP295 (Spray, Oregon), MSH22, MSH38 (Mt. St. Helena, California). Flies were obtained by S. Yi from the National *Drosophila* Species Resource Center (Bowling Green, Ohio), M. Noor and W. Anderson, and supplied to us. The strains were cultured on banana medium at 18°C.

Microsatellite loci

Noor *et al.* recently identified and mapped microsatellites from *D. pseudoobscura*, a sister species of *D. miranda* [48]. Based on Muller's chromosome arm homology [10], we selected nine microsatellites located on the X chromosome of *D. miranda,* nine loci from chromosome 2, and nine loci located on the neo-X/neo-Y chromosome (see Tables 1 and 2). For the X-linked and the autosomal microsatellites, a single male fly of each line was analysed. To unambiguously distinguish between the neo-X and neo-Y allele of a microsatellite locus, single male–female crosses were performed, and the genotypes of the parents and five F_1 males were determined. Single-fly DNA was extracted by squash preparations [49].

Characterisation of microsatellites

Microsatellite typing was performed by PCR amplification with direct incorporation of [α-^{32}P]dCTP. PCR reactions were carried out in 10 µl reaction volumes (2 mM $MgCl_2$, 1 mM dGTP, dATP and dTTP, 0.1 mM dCTP, 0.4 µM of each primer, 50–100 ng template DNA and 0.1 µCi [α-^{32}P]dCTP), using the same annealing temperature for each locus as in [48]. The PCR products were separated on a sequencing gel and visualised by autoradiography. Alleles were sized by running an M13 control sequence on each gel.

Analyses of microsatellite data

Genetic variation was quantified by determining the number of alleles at each locus and by calculating the variance in repeat number and the heterozygosity. The variance in repeat number was calculated as $V = n\Sigma p_i (x_i - \bar{x})^2/(n-1)$ and heterozygosity as $H = n(1 - \Sigma p_i^2)/(n-1)$ [50], where n is the number of chromosomes sampled, p_i is the frequency of the ith allele, x_i is the number of repeats at the ith allele and $\bar{x}$ is the average number of repeats. In the few cases where a heterozygous individual was scored at an autosomal locus, one allele was randomly discarded to account for inbreeding in small laboratory populations.

Population differentiation

Two different approaches were used to investigate whether the American and the Canadian populations show genetic differentiation. The first method tests whether the distribution of alleles in the various populations is identical across populations. For each locus, an exact contingency-table test of population subdivision was performed, using the GENEPOP package [51]. The second approach involved the calculation of F_{ST} [52]. The significance level for F_{ST} was calculated by randomly permuting individuals among population and determining the proportion of resampled individuals with a F_{ST} equal or larger than the one observed between the two populations, using the software package GENETIX [53].

Coalescence simulations

Coalescence simulations using the standard algorithm of Hudson [54] were performed to determine the value of N_e of the neo-Y compared with the neo-X chromosome. While all loci on the neo-Y chromosome share one genealogy, the microsatellites on the neo-X chromosome are probably completely independent (they show no evidence for linkage disequilibrium in our data; results not shown). Therefore, we generated seven independent trees for a data set of 12 chromosomes for the neo-X chromosome, whereas a single tree with seven completely linked loci was computed for the neo-Y chromosome. Mutations were superimposed on the trees following the single stepwise mutation model [55]. For each run, seven random estimates of $\theta = 4N_e\mu$ were drawn from a gamma prior distribution with parameters (0.55, 10) which has a mean of 5.5 and a variance of 55. These values were chosen because they gave a good fit to the estimated variance in repeat number for microsatellites located on the X chromosome and autosomes of *D. miranda* ($\overline{V} = 5.3$ and $\mathrm{Var}(V) = 58$). Microsatellite mutations using the respective θ values were laid down on the trees according to Poisson distributions, using the same values of θ for homologous loci. The average variance in repeat number $\overline{V}$ per locus across the seven loci was calculated. After each run, $\Delta V = (\overline{V}_X - \overline{V}_Y)/\overline{V}_X$ for the simulated data set was computed and compared to the observed value of ΔV. This procedure was repeated 10^6 times and the fraction of simulated trees, P_c, with a ΔV equal to or larger than the observed ΔV was reported. To estimate confidence intervals for the reduction in N_e of the neo-Y loci, the values of θ were multiplied by a scaling factor k before mutations were laid down on the neo-Y tree.

Sequencing of neo-X and neo-Y alleles

To verify that a microsatellite is present on the neo-sex chromosomes in *D. miranda*, we sequenced one neo-X and one neo-Y allele from each locus. To obtain a template for sequencing, a PCR reaction was performed using the genomic DNA of a single male of known genotype, followed by separation of the products on a 2.5% agarose gel and extraction of the DNA using the Qiagen gel extraction kit. The extracted PCR fragments were cloned into a derivative of the pZero2.1 vector (Invitrogen). Both strands were sequenced on an ABI377 automated sequencer using the ABI PRISM BigDye termination cycle-sequencing kit (Perkin-Elmer). The total number of nucleotides determined was 938 for the neo-X and 914 for the neo-Y (excluding the primer sequence). Mutations in the microsatellite structure and insertion/deletion polymorphism were excluded for the purpose of further analyses, resulting in a total of 733 bp of flanking sequence. Genetic distances using the combined sequences from all loci were calculated using the method of Jukes and Cantor [56]. Potential PCR errors (~1 error per kb sequence) would reduce our divergence estimates slightly (from 0.025 to ~0.023).

Acknowledgements

We are very grateful to J. Pemberton for lab space, and to J. Pritchard for providing us with his simulation algorithm. B.C. is supported by the Royal Society and D.B. by the Austrian Science Ministry. S. Yi provided us with the strains of *D. miranda*. We thank S. Yi and two reviewers for their comments on the manuscript.

References

1. Bull JJ: *Evolution of Sex Determining Mechanisms*. Menlo Park: Benjamin Cummings; 1983.
2. Westergaard M: **The mechanism of sex determination in flowering plants.** *Adv Genet* 1958, **9:**217-281.
3. Charlesworth B: **The evolution of chromosomal sex determination and dosage compensation.** *Curr Biol* 1996, **6:**149-162.
4. Lucchesi JC: **On the origin of sex chromosomes.** *Bioessays* 1999, **21:**188-190.
5. Rice WR: **Evolution of the Y sex chromosome in animals.** *Bioscience* 1996, **46:**331-343.
6. Charlesworth B: **Model for evolution of Y chromosomes and dosage compensation.** *Proc Natl Acad Sci USA* 1978, **75:**5618-5622.
7. McVean GAT, Charlesworth B: **The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation.** *Genetics* 2000, **155**:929-944.
8. Rice WR: **Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome.** *Genetics* 1987, **116:**161-167.
9. Steinemann M, Steinemann S: **Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution.** *Genetica* 1998, **102-103:**409-420.
10. Muller HJ: **Bearings of the *Drosophila* work on systematics.** In *The New Systematics*. Edited by Huxley JS. Oxford: Oxford University Press; 1940:185-268.
11. Ashburner M: *Drosophila: A Laboratory Handbook*. New York: Cold Spring Harbor Laboratory Press; 1989.
12. Steinemann M, Steinemann S, Lottspeich F: **How Y chromosomes become genetically inert.** *Proc Natl Acad Sci USA* 1993, **90:**5737-5741.
13. Steinemann S, Steinemann M: **The *Amylase* gene cluster on the evolving sex chromosomes of *Drosophila miranda*.** *Genetics* 1999, **151:**151-161.
14. Marin I, Franke A, Bashaw GJ, Baker BS: **The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes.** *Nature* 1996, **383:**160-163.
15. Bone JR, Kuroda MI: **Dosage compensation regulatory proteins and the evolution of sex chromosomes in *Drosophila*.** *Genetics* 1996, **144:**705-713.
16. Steinemann M, Steinemann S, Turner BM: **Evolution of dosage compensation.** *Chromosome Res* 1996, **4:**1-6.
17. Kimura M: *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press; 1983.
18. Yi S, Charlesworth B: **Contrasting patterns of molecular evolution of the genes on the new and old sex chromosomes of *Drosophila miranda*.** *Mol Biol Evol* 2000, **17**:703-717.
19. Filatov DA, Moneger F, Negrutiu I, Charlesworth D: **Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution.** *Nature* 2000, **404:**388-390.
20. Moran PAP: **Wandering distributions and the electrophoretic profile.** *Theor Popul Biol* 1975, **8:**318-330.
21. Yi S: **Molecular variation and the evolution of the neo-sex chromosomes of *Drosophila miranda*.** PhD Dissertation. Chicago: University of Chicago; 2000.
22. Wright S: *Evolution and the Genetics of Populations. Vol.2 The Theory of Gene Frequencies*. Chicago: University of Chicago Press; 1969.
23. Wang RL, Hey J: **The speciation history of *Drosophila pseudoobscura* and close relatives: inferences from DNA sequence variation at the *period* locus.** *Genetics* 1996, **144:**1113-1126.
24. Roewer L, Arnemann J, Spurr NK, Grzeschik KH, Epplen JT: **Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts.** *Hum Genet* 1992, **89:**389-394.
25. Perez-Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E, *et al.*: **Population genetics of Y-chromosome short tandem repeats in humans.** *J Mol Evol* 1997, **45:**265-270.
26. Goldstein DB, Zhivotovsky LA, Nayar K, Linares AR, Cavalli-Sforza LL, Feldman MW: **Statistical properties of the variation at linked microsatellite loci: implications for the history of human Y chromosomes.** *Mol Biol Evol* 1996, **13:**1213-1218.
27. Lahn BT, Page DC: **Four evolutionary strata on the human X chromosome.** *Science* 1999, **286:**964-967.
28. Lahn BT, Page DC: **Functional coherence of the human Y chromosome.** *Science* 1997, **278:**675-680.
29. Zurovcova M, Eanes WF: **Lack of nucleotide polymorphism in the Y-linked sperm flagellar dynein gene *Dhc-Yh3* of *Drosophila melanogaster* and *D. simulans*.** *Genetics* 1999, **153:**1709-1715.
30. McAllister BF, Charlesworth B: **Reduced sequence variability on the neo-Y chromosome of *Drosophila americana americana*.** *Genetics* 1999, **153:**221-233.
31. Charlesworth B, Charlesworth D, Hnilicka J, Yu A, Guttman DS: **Lack of degeneration of loci on the neo-Y chromosome of *Drosophila americana americana*.** *Genetics* 1997, **145:**989-1002.
32. Henderson ST, Petes TD: **Instability of simple sequence DNA in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1992, **12:**2749-2757.
33. Levinson G, Gutman GA: **Slipped-strand mispairing: a major mechanism for DNA sequence evolution.** *Mol Biol Evol* 1987, **4:**203-221.
34. Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P: **Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees.** *Hum Mol Genet* 1997, **6:**799-803.

35. Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, *et al.*: **Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome as revealed by direct observation in father/son pairs.** *Am J Hum Genet* 2000, **66:**1580-1588.
36. Schlötterer C, Ritter R, Harr B, Brem G: **High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates.** *Mol Biol Evol* 1998, **15:**1269-1274.
37. Caballero A: **On the effective size of populations with separate sexes with particular reference to sex-linked genes.** *Genetics* 1995, **139:**1007-1011.
38. Nunney L: **The influence of mating system and overlapping generations on effective population size.** *Evolution* 1993, **47:**1329-1341.
39. Maynard Smith J, Haigh J: **The hitch-hiking effect of a favourable gene.** *Genet Res* 1974, **23:**23-35.
40. Kaplan NL, Hudson RR, Langley CH: **The 'hitchhiking effect' revisited.** *Genetics* 1989, **123:**887-899.
41. Charlesworth B, Morgan MT, Charlesworth D: **The effect of deleterious mutations on neutral molecular variation.** *Genetics* 1993, **134:**1289-1303.
42. Charlesworth B: **The effect of background selection against deleterious mutations on weakly selected linked variants.** *Genet Res* 1994, **63:**213-227.
43. Hudson RR, Kaplan NL: **Deleterious background selection with recombination.** *Genetics* 1995, **141:**1605-1617.
44. Muller H: **The relation of recombination to mutational advance.** *Mut Res* 1964, **1:**2-9.
45. Gordo I, Charlesworth B: **The degeneration of asexual haploid populations and the speed of Muller's ratchet.** *Genetics* 2000, **154:**1379-1387.
46. Charlesworth B, Charlesworth D: **Rapid fixation of deleterious alleles can be caused by Muller's ratchet.** *Genet Res* 1997, **70:**63-73.
47. Hill WG, Robertson A: **The effect of linkage on limits to artificial selection.** *Genet Res* 1966, **8:**269-294.
48. Noor MAF, Schug MD, Aquadro CF: **Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis.*** *Genet Res* 2000, **75:**25-35.
49. Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK, *et al.*: **Type I repressors of P element mobility.** *Genetics* 1993, **135:**81-95.
50. Nei M: **Genetic distance between populations.** *Am Nat* 1972, **106:**283-292.
51. Raymond M, Rousset F: **GENEPOP: population genetics software for exact tests and ecumenicism.** *J Heredity* 1995, **86:**248-249.
52. Weir BS, Cockerham CC: **Estimating F-statistics for the analysis of population structure.** *Evolution* 1984, **38:**1358-1370.
53. Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F: **GENETIX.** Montpellier; 1996-98, http://wwwuniv-montp2fr/~genetix/genetixhtm.
54. Hudson RR: **Gene genealogies and the coalescent process**. In *Oxford Surveys in Evolutionary Biology.* Edited by Antonovics J, Futuyma D. Oxford: Oxford University Press; 1990: 1-44.
55. Ohta T, Kimura M: **A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population.** *Genet Res* 1973, **22:**201-204.
56. Jukes TH, Cantor CR: **Evolution of protein molecules**. In *Mammalian Protein Metabolism.* Edited by Munro HN. New York: Academic Press; 1969:21-123.

Because *Current Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cub – for further information, see the explanation on the contents page.